

Datenbereinigung mit OpenRefine



OpenRefine

Jakob Frohmann
j.frohmann@ub.uni-frankfurt.de

Vorbereitung | Hinweise

- Melden Sie sich für den moodle-Kurs [Praxislabor Digitale Geisteswissenschaften](http://www.ub.uni-frankfurt.de/digitalhumanities) an und schreiben Sie sich ein:



**Praxislabor Digitale Geisteswissenschaften:
Einführungsworkshops zu Methoden und
Werkzeugen der Digital Humanities**

Praxislabor Digitale Geisteswissenschaften:
Einführungsworkshops zu Methoden und Werkzeugen der Digital Humanities

Zum ersten Kennenlernen von Methoden und Werkzeugen der Digital Humanities bietet die Universitätsbibliothek JCS (im Bibliothekszentrum Geisteswissenschaften) Studierenden und Mitarbeiter/-innen der Goethe-Uni im kommenden Wintersemester Workshops an. In niederschwelligen Einführungen werden anhand von überschaubaren, konkreten Beispielen aus der Praxis Methoden, Tools oder Themen der digitalen Geisteswissenschaften vorgestellt und geübt und so ein erster Einblick in die Möglichkeiten gegeben, wie klassische Methoden der Geisteswissenschaften mithilfe digitaler Verfahren der Textanalyse sowie der Text- und Datenaufbereitung sinnvoll ergänzt werden können.

Die Workshopreihe besteht jeweils aus inhaltlich zusammenhängenden Zweierblöcken, in denen auf eine Präsentation eine Sitzung zur Vertiefung und Übung folgt.

Die Workshops richten sich an interessierte Einsteiger; besondere Kenntnisse werden nicht vorausgesetzt. Nähere Informationen sowie die Möglichkeit zur Anmeldung finden Sie unter: <http://www.ub.uni-frankfurt.de/digitalhumanities/workshops.html>.

Praxislabor Digitale Geisteswissenschaften: Einführungsworkshops zu Methoden und Werkzeugen der Digital Humanities

Startseite / Kurse / Verschiedenes / Praxislabor Digitale Geisteswissenschaften

Allgemeines

In dieses kollaborative Dokument können Themenvorschläge für die Hands-on-Sessions eingetragen werden.

Informationen zur Anmeldung und Kurszeiten unter: <http://www.ub.uni-frankfurt.de/digitalhumanities>

Einführung in TEI / XML

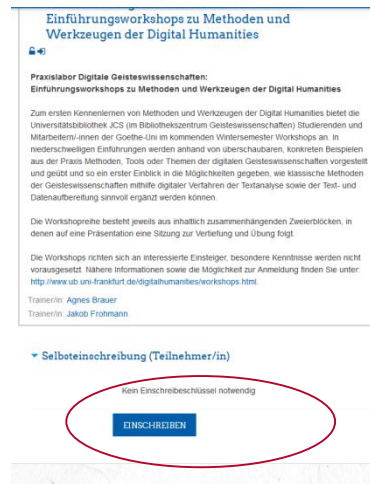
Dozentin: Agnes Brauer

Der Workshop führt in die Grundlagen der Textauszeichnung mit TEI ein, einer XML-basierten und sich mittlerweile als De-facto-Standard etablierten Auszeichnungssprache speziell für die Zwecke der Geisteswissenschaften. Nach einer knappen allgemeinen Einführung werden die Teilnehmer anhand einer kleinen Übung die Praxis der Textauszeichnung mit TEI kennenlernen und sich einen ersten Überblick über die Bedeutung und die verschiedenen Module dieser Sprache verschaffen.

Link: <http://www.tei-c.org/>

Hands-on Übung zur TEI/XML-Einführung

Dozentin: Agnes Brauer



**Einführungsworkshops zu Methoden und
Werkzeugen der Digital Humanities**

Praxislabor Digitale Geisteswissenschaften:
Einführungsworkshops zu Methoden und Werkzeugen der Digital Humanities

Zum ersten Kennenlernen von Methoden und Werkzeugen der Digital Humanities bietet die Universitätsbibliothek JCS (im Bibliothekszentrum Geisteswissenschaften) Studierenden und Mitarbeiter/-innen der Goethe-Uni im kommenden Wintersemester Workshops an. In niederschwelligen Einführungen werden anhand von überschaubaren, konkreten Beispielen aus der Praxis Methoden, Tools oder Themen der digitalen Geisteswissenschaften vorgestellt und geübt und so ein erster Einblick in die Möglichkeiten gegeben, wie klassische Methoden der Geisteswissenschaften mithilfe digitaler Verfahren der Textanalyse sowie der Text- und Datenaufbereitung sinnvoll ergänzt werden können.

Die Workshopreihe besteht jeweils aus inhaltlich zusammenhängenden Zweierblöcken, in denen auf eine Präsentation eine Sitzung zur Vertiefung und Übung folgt.

Die Workshops richten sich an interessierte Einsteiger; besondere Kenntnisse werden nicht vorausgesetzt. Nähere Informationen sowie die Möglichkeit zur Anmeldung finden Sie unter: <http://www.ub.uni-frankfurt.de/digitalhumanities/workshops.html>

Trainerin: Agnes Brauer
Trainerin: Jakob Frohmann

▼ **Selbsteinschreibung (Teilnehmer/in)**

Kein Einschreibeschlüssel notwendig

EINSCHREIBEN

Vorbereitung | Hinweise

- Laden Sie bitte OpenRefine herunter (93MB) und installieren / entpacken Sie die Software auf Ihrem Computer. Sie benötigen den Browser Firefox, in dem OpenRefine läuft (Java-Programm).

<http://openrefine.org/download.html>

<https://github.com/OpenRefine/OpenRefine/releases>

- Starten Sie die Anwendung aus dem entpackten Verzeichnis, es öffnen sich eine Kommandozeile und kurz danach der Browser mit dem geladenen Programm – sollte der Browser nicht starten, benutzen Sie bitte den Link:

<http://127.0.0.1:3333/>.

- Tragen Sie bitte bei Bedarf / Interesse Themenvorschläge für die Hands-on Übung in das [kollaborative Dokument](#) ein.

OpenRefine – Was ist das?

- Interaktives Werkzeug zum Bearbeiten, Erkunden und Bereinigen großer Mengen von Daten in Tabellenform („A power tool for working with messy data.“)
- Hat große Ähnlichkeiten zu einem Tabellenkalkulationsprogramm mit Zeilen und Spalten (z.B. Excel), aber funktioniert eher wie eine relationale Datenbank.
(ein „OpenRefine-Projekt“ = eine Tabelle)
- Läuft lokal auf dem Rechner, aber im Browser (keine Internetverbindung notwendig)
- Open Source-Software in Java
(zuvor zwischenzeitlich zu Google gehörig unter dem Namen Google Refine)

Anwendungsmöglichkeiten und Basisfunktionen

- Reinigung und Aufbereitung von Daten
- „Exploration“ von Daten, Aufspüren von Inkonsistenzen oder Fehlern im Datenformat („the big picture of your data“)
- Umstrukturierung und Überführung von Daten in eine andere Form

Anwendungsmöglichkeiten und Basisfunktionen

- wichtige Werkzeuge in OpenRefine: verschiedene Arten von Filtern und Facetten mit vordefinierten Kriterien zum Anpassen der Anzeige
- Alle Operationen geschehen i.d.R. nur auf den gewählten / selektierten Daten.
- Alle Veränderung der Daten geschehen auf einer Kopie des Datensets und können leicht wieder rückgängig gemacht werden ("play with your data").
- Abfolgen von Operationen können gespeichert und dann auf andere Datensätze ebenfalls angewendet werden.

Anwendungsmöglichkeiten – fortgeschrittene Funktionen

Reconcile & Match

- Vergleichen / Angleichen der eigenen Daten anhand von Datenbanken (z.B. Wikidata)
- Anreicherung von Daten (z.B. mit eindeutigen Identifikatoren)
- Verlinkung von Daten

... mit Hilfe von diversen Webservices mit offenen Schnittstellen (APIs) über das Internet.

Grundsätzliche Arbeitsweise mit Open Refine

Die Daten herausfilter, welche man verändern möchte, dann die selektierten Daten gemeinsam in einer Operation bearbeiten.

Vorgehensweise:

- Quelldaten einlesen
- Daten analysieren
- Daten aufräumen und optimieren
- Daten anreichern
- Daten im Zielformat ausgeben

Siehe auch: <https://www.tanner.de/blog/systeme-technologien/datenanalyse-und-optimierung-mit-openrefine-1/>

Daten importieren/exportieren aus OpenRefine

- mögliche Dateiformate für den Import (Auswahl): TSV, CSV, Microsoft Excel, JSON, XML
- mögliche Dateiformate für den Export (Auswahl): TSV, CSV, Microsoft Excel, HTML (Tabelle)

Bibl. Daten DNB - Stichwort Nationalsozialismus, 500 Treffer - Excel

Start Einfügen Seitenlayout Formeln Daten Überprüfen Ansicht Was möchten Sie tun?

AutoSumme Füllbereich Löschen Sortieren und Filtern Suchen und Auswählen

ISBN/Einband/Preis: Broschur

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1																						
2	Logo der Deutschen Nationalbibliothek																					
3	Schriftbanner mit Leipzig, Frankfurt am Main																					
4	Katalog der Deutschen Nationalbibliothek																					
5	Suchergebnis - Suche nach: "Nationalsozialismus", Treffer 1-500 von insgesamt 13700 - 16.01.19 11:11																					
6	Bücher	1																				
7	Titel: Kunst am Pranger : Die Moderne im Nationalsozialismus / Andreas Hüneke																					
8	Person(en): Hüneke, Andreas																					
9	Ausgabe: 2018. Auflage																					
10	Verlag: Paderborn : Fink, Wilhelm																					
11	Zeitliche Einordnung: Erscheinungstermin: Dezember 2021																					
12	Umfang/Format: 280 Seiten : 40 s/w Abb.																					
13	ISBN/Einband/Preis: 978-3-7705-5183-5 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen] Gb. : ca. EUR 34.90 (DE), ca. sfr 49.90 (freier Pr.)																					
14	Bestellnummer(n): VLB_ID_old: 3633739																					
15	EAN: 9783770551835 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen]																					
16	Sprache(n): Deutsch (ger)																					
17	Beziehungen: Schriften der Berliner Forschungsstelle „Entartete Kunst“																					
18	Sachgruppe(n): 700 Künste, Bildende Kunst allgemein																					
19	Bücher	2																				
20	Titel: Nationalsozialismus und Exil (1933-1945) : Lehrbuch Germanistik / Wilhelm Haefs																					
21	Person(en): Haefs, Wilhelm																					
22	Organisation(en): J.-B.-Metzlersche Verlagsbuchhandlung und Carl-Ernst-Poeschel-Verlag (Verlag)																					
23	Ausgabe: 1. Auflage 2021																					
24	Verlag: Stuttgart : J.B. Metzler, Part of Springer Nature - Springer-Verlag GmbH - J.B. Metzler																					
25	Zeitliche Einordnung: Erscheinungstermin: Juli 2021																					
26	Umfang/Format: 250 Seiten in 1 Teil ; 23.5 cm x 15.5 cm																					
27	Andere Ausgabe(n): Erscheint auch als Online-Ausgabe: ISBN: 9783476054661 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen]																					
28	ISBN/Einband/Preis: 978-3-476-02585-2 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen] Festeinband : EUR 19.95 (DE), EUR 20.51 (AT), CHF 22.00 (freier Preis); 3-476-02585-3 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen]																					

Tabelle1

Bearbeitungsbeispiele

Data you have	Desired data
1st January 2014	2014-01-01
01/01/2014	2014-01-01
Jan 1 2014	2014-01-01
2014-01-01	2014-01-01

<https://librarycarpentry.org/lc-open-refine/03-working-with-data/index.html>

Bearbeitungsbeispiele

Data you have	Desired data
London	London
London]	London
London,]	London
london	London

<https://librarycarpentry.org/lc-open-refine/03-working-with-data/index.html>

Bearbeitungsbeispiele

Address in single field	Institution	Library name	Address 1	Address 2	Town/City	Region	Country	Postcode
University of Wales, Llyfrgell Thomas Parry Library, Llanbadarn Fawr, ABERYSTWYTH, Ceredigion, SY23 3AS, United Kingdom	University of Wales	Llyfrgell Thomas Parry Library	Llanbadarn Fawr		Aberystwyth	Ceredigion	United Kingdom	SY23 3AS
University of Aberdeen, Queen Mother Library, Meston Walk, ABERDEEN, AB24 3UE, United Kingdom	University of Aberdeen	Queen Mother Library	Meston Walk		Aberdeen		United Kingdom	AB24 3UE
University of Birmingham, Barnes Library, Medical School, Edgbaston, BIRMINGHAM, West Midlands, B15 2TT, United Kingdom	University of Birmingham	Barnes Library	Medical School	Edgbaston	Birmingham	West Midlands	United Kingdom	B15 2TT
University of Warwick, Library, Gibbett Hill Road, COVENTRY, CV4 7AL, United Kingdom	University of Warwick	Library	Gibbett Hill Road		Coventry		United Kingdom	CV4 7AL

<https://librarycarpentry.org/lc-open-refine/03-working-with-data/index.html>


Ein Projekt erstellen...

← → ↺ 🏠

127.0.0.1:3333

⋮ 📌 ☆

⚙ Meistbesucht 🌈 Erste Schritte

 **OpenRefine** *A power tool for working with messy data.*

Create Project

Open Project

Import Project

Language Settings

Create a project by importing data. What kinds of data files can I import?

TSV, CSV, *SV, Excel (.xls and .xlsx), JSON, XML, RDF as XML, and Google Data documents are all supported. Support for other formats can be added with OpenRefine extensions.

Get data from

This Computer

Web Addresses (URLs)

Clipboard

Data Package (JSON URL)

Database


Google Data

Locate one or more files on your computer to upload:

Durchsuchen...

Keine Dateien ausgewählt.

Next »



Version 3.1 [b90e413]

Preferences

Help

About

Ein Projekt erstellen...

← → ↺ ↻ 127.0.0.1:3333 ... ⌵ ⭐

Meistbesucht Erste Schritte

OpenRefine A power tool for working with messy data.

Create Project « Start Over Configure Parsing Options Project name Tags **Create Project »**

Open Project
Import Project
Language Settings

Column 1	Column 2	Column 3	Column 4	Column 5
1. Logo der Deutschen Nationalbibliothek				
2. Schriftbanner mit Leipzig, Frankfurt am Main				
3. Katalog der Deutschen Nationalbibliothek				
4. Suchergebnis - Suche nach: "Nationalsozialismus", Treffer 1-500 von insgesamt 13700 - 16.01.19 11:11				
5. Bücher		1		Link zu diesem Datensatz: http://d-nb.info/1009008528 [http://d-nb.info/1009008528]
6. Titel: Kunst am Pranger : Die Moderne im Nationalsozialismus / Andreas Hüneke				
7. Person(en): Hüneke, Andreas				
8. Ausgabe: 2018. Auflage				
9. Verlag: Paderborn : Fink, Wilhelm				
10. Zeitliche Einordnung: Erscheinungstermin: Dezember 2021				
11. Umfang/Format: 280 Seiten : 40 s/w Abb.				
12. ISBN/Einband/Preis: 978-3-7705-5183-5 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen] Gb. : ca. EUR 34.90 (DE), ca. sfr 49.90 (freier Pr.)				
13. Bestellnummer(n): VLB_ID_old: 3633739				
14. EAN: 9783770551835 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen]				
15. Sprache(n): Deutsch (ger)				

Parse data as

Excel files

JSON files

Line-based text files

CSV / TSV / separator-based files

Fixed-width field text files

PC-Axis text files

MARC files

JSON-LD files

RDF/N3 files

RDF/N-Triples files

Worksheets to Import

☒ Bibl. Daten DNB - Stichwort Nationalsozialismus, 500 Treffer.xlsx#Tabelle1 7160 rows

☐ Ignore first 0 line(s) at beginning of file

☒ Parse next 1 line(s) as column headers

☐ Discard initial 0 row(s) of data

☐ Load at most 0 row(s) of data

☒ Store blank rows

☒ Store blank cells as nulls

☐ Store file source (file names, URLs) in each row

Update Preview

Version 3.1 [b90e413]

Preferences
Help
About

Ein Projekt erstellen...

127.0.0.1:3333/project?project=1988236490708

Meistbesucht Erste Schritte

OpenRefine Bibl Daten DNB Stichwort Nationalsozialismus 500 Treffer.xlsx Permalink

Open... Export Help

Facet / Filter Undo / Redo 0 / 0

7160 rows

Show as: rows records Show: 5 10 25 50 rows

« first < previous 1 - 50 next > last »

Extensions: Wikidata

Using facets and filters

Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top of each data column.

Not sure how to get started?
[Watch these screencasts](#)

All	Column 1	Column 2	Column 3	Column 4	Column 5
3.	Katalog der Deutschen Nationalbibliothek				
4.	Suchergebnis - Suche nach: "Nationalsozialismus", Treffer 1-500 von insgesamt 13700 - 16.01.19 11:11				
5.	Bücher		1		Link zu diesem Datensatz http://d-nb.info/1009008 [http://d-nb.info/1009008]
6.	Titel: Kunst am Pranger : Die Moderne im Nationalsozialismus / Andreas Hüneke				
7.	Person(en): Hüneke, Andreas				
8.	Ausgabe: 2018. Auflage				
9.	Verlag: Paderborn : Fink, Wilhelm				
10.	Zeitliche Einordnung: Erscheinungstermin: Dezember 2021				
11.	Umfang/Format: 280 Seiten : 40 s/w Abb.				
12.	ISBN/Einband/Preis: 978-3-7705-5183-5 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen] Gb. : ca. EUR 34.90 (DE), ca. sfr 49.90 (freier Pr.)				
13.	Bestellnummer(n): VLB_ID_old: 3633739				
14.	EAN: 9783770551835 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen]				
15.	Sprache(n): Deutsch (ger)				
16.	Beziehungen: Schriften der Berliner Forschungsstelle „Entartete Kunst“				
17.	Sachgruppe(n): 700 Künste, Bildende Kunst allgemein				
18.	Bücher		2		Link zu diesem Datensatz http://d-nb.info/1173826- [http://d-nb.info/1173826-]
19.	Titel: Nationalsozialismus und Exil (1933-1945) : Lehrbuch Germanistik / Wilhelm Haefs				

Ein Projekt erstellen...

← → ↺ 🏠 127.0.0.1:3333 ... 📌 ☆

⚙ Meistbesucht 🚦 Erste Schritte

OpenRefine *A power tool for working with messy data.*

Create Project « Start Over Configure Parsing Options Project name Tags **Create Project »**

Open Project
Import Project
Language Settings

Column 1

1. 1»
2. Logo der Deutschen Nationalbibliothek
3. Schriftbanner mit Leipzig, Frankfurt am Main
4. Katalog der Deutschen Nationalbibliothek
5. Suchergebnis - Suche nach: "Nationalsozialismus", Treffer 1-500 von insgesamt 13700 - 16.01.19 11:11
6. BÄcher 1 Link zu diesem Datensatz: <http://d-nb.info/1009008528> [http://d-nb.info/1009008528]
7. Titel: Kunst am Pranger : Die Moderne im Nationalsozialismus / Andreas Häneke
8. Person(en): Häneke, Andreas
9. Ausgabe: 2018. Auflage
10. Verlag: Paderborn : Fink, Wilhelm
11. Zeitliche Einordnung: Erscheinungstermin: Dezember 2021
12. Umfang/Format: 280 Seiten : 40 s/w Abb.
13. ISBN/Einband/Preis: 978-3-7705-5183-5 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen] Gb. : ca. EUR 34.90 (DE), ca. sfr 49.90 (freier Pr.)
14. Bestellnummer(n): VLB_ID_old: 3633739
15. EAN: 9783770551835 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen]
16. Sprache(n): Deutsch (ger)
17. Beschreibung: Schöner des Radlers Erscheinungstermin: 2018

Parse data as

Line-based text files

CSV / TSV / separator-based files

Fixed-width field text files

PC-Axis text files

JSON files

MARC files

JSON-LD files

RDF/N3 files

RDF/N-Triples files

RDF/Turtle files

Character encoding

Parse every 1 lines into one row

☒ Store blank rows

☒ Store blank cells as nulls

☐ Store file source (file names, URLs) in each row

☐ Ignore first 0 line(s) at beginning of file

☐ Discard initial 0 row(s) of data

☐ Load at most 0 row(s) of data

Update Preview

Version 3.1 [b90e413]

Preferences
Help
About

Ein Projekt erstellen...

OpenRefine A power tool for working with messy data.

Meistbesucht Erste Schritte

127.0.0.1:3333

« Start Over Configure Parsing Options Project name Bibl Daten DNB Stichwort National Tags Create Project »

	Column 2	Column 3	Column 4	Column 5
1. Logo der Deutschen Nationalbibliothek				
2. Schriftbanner mit Leipzig	Frankfurt am Main			
3. Katalog der Deutschen Nationalbibliothek				
4. Suchergebnis - Suche nach: "Nationalsozialismus, Treffer 1-500 von insgesamt 13700 - 16.01.19 11:11 Bücher 1 Link zu diesem Datensatz: http://d-nb.info/1009008528 [http://d-nb.info/1009008528] Titel: Kunst am Pranger : Die Moderne im Nationalsozialismus / Andreas Hüneke Person(en): Hüneke, Andreas Ausgabe: 2018. Auflage Verlag: Paderborn : Fink, Wilhelm Zeitliche Einordnung: Erscheinungstermin: Dezember 2021 Umfang/Format: 280 Seiten : 40 s/w Abb. ISBN/Einband/Preis: 978-3-7705-5183-5 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen] Gb. : ca. EUR 34.90 (DE), ca. sfr 49.90 (freier Pr.) Bestellnummer(n): VLB_ID_old: 3633739 EAN: 9783770551835 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen] Sprache(n): Deutsch (ger) Beziehungen: Schriften der Berliner Forschungsstelle „Entartete Kunst“ Sachgruppe(n): 700 Künste, Bildende Kunst allgemein Bücher 2 Link zu diesem Datensatz: http://d-nb.info/1173826467 [http://d-nb.info/1173826467] Titel: Nationalsozialismus und Exil (1933-1945) : Lehrbuch Germanistik / Wilhelm Haefs Person(en): Haefs, Wilhelm				

Parse data as

Line-based text files

CSV / TSV / separator-based files

Fixed-width field text files

PC-Axis text files

JSON files

MARC files

JSON-LD files

RDF/N3 files

RDF/N-Triples files

RDF/Turtle files

Character encoding UTF-8

Columns are separated by

☒ commas (CSV)

☐ tabs (TSV)

☐ custom: \t

Escape special characters with \

☐ Column names (comma separated):

☐ Ignore first 0 line(s) at beginning of file

☒ Parse next 1 line(s) as column headers

☐ Discard initial 0 row(s) of data

☐ Load at most 0 row(s) of data

☒ Use character " to enclose cells containing column separators

☐ Parse cell text into numbers, dates, ...

☒ Store blank rows

☒ Store blank cells as nulls

☐ Store file source (file names, URLs) in each row

Update Preview

Version 3.1 [b90e413]

Preferences

Help

About

Ein Projekt erstellen...

← → ↺ 🏠 127.0.0.1:3333 ... ☆

⚙ Meistbesucht 📁 Erste Schritte

OpenRefine *A power tool for working with messy data.*

Create Project « Start Over Configure Parsing Options Project name Tags **Create Project »**

Open Project

Import Project

Language Settings

	Column	Column 3	Column 4	Column 5
1. Logo der Deutschen Nationalbibliothek				
2. Schriftbanner mit Leipzig, Frankfurt am Main				
3. Katalog der Deutschen Nationalbibliothek				
4. Suchergebnis - Suche nach: "Nationalsozialismus", Treffer 1-500 von insgesamt 13700 - 16.01.19 11:11				
5. Bücher		1		Link zu diesem Datensatz: http://d-nb.info/1009008528 [http://d-nb.info/1009008528]
6. Titel: Kunst am Pranger : Die Moderne im Nationalsozialismus / Andreas Hüneke				
7. Person(en): Hüneke, Andreas				
8. Ausgabe: 2018. Auflage				
9. Verlag: Paderborn : Fink, Wilhelm				
10. Zeitliche Einordnung: Erscheinungstermin: Dezember 2021				
11. Umfang/Format: 280 Seiten : 40 s/w Abb.				
12. ISBN/Einband/Preis: 978-3-7705-5183-5 ⓘ [Titel anhand dieser ISBN in Citavi-Projekt übernehmen] Gb. : ca. EUR 34.90 (DE), ca. sfr 49.90 (freier Pr.)				
13. Bestellnummer(n): VLB_ID_old: 3633739				
14. EAN: 9783770551835 ⓘ [Titel anhand dieser ISBN in Citavi-Projekt übernehmen]				
15. Sprache(n): Deutsch (ger)				

Parse data as

Line-based text files

CSV / TSV / separator-based files

Fixed-width field text files

PC-Axis text files

JSON files

MARC files

JSON-LD files

RDF/N3 files

RDF/N-Triples files

RDF/Turtle files

Character encoding

Columns are separated by

☐ commas (CSV)

☒ tabs (TSV)

☐ custom: \t

Escape special characters with \

☐ Column names (comma separated):

☐ Ignore first 0 line(s) at beginning of file

☒ Parse next 1 line(s) as column headers

☐ Discard initial 0 row(s) of data

☐ Load at most 0 row(s) of data

☒ Use character " to enclose cells containing column separators

☐ Parse cell text into numbers, dates, ...

☒ Store blank rows

☒ Store blank cells as nulls

☐ Store file source (file names, URLs) in each row

Update Preview

Version 3.1 [b90e413]

Preferences

Help

About

javascript:()

Ein Projekt erstellen...

127.0.0.1:3333/project?project=1893534342500

Meistbesucht Erste Schritte

OpenRefine Bibl Daten DNB Stichwort Nationalsozialismus 500 Treffer txt [Permalink](#) [Open...](#) [Export](#) [Help](#)

Facet / Filter Undo / Redo 0 / 0

5093 rows Extensions: Wikidata

Show as: rows records Show: 5 10 25 50 rows « first < previous 1 - 50 next > last »

Using facets and filters

Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top of each data column.

Not sure how to get started? [Watch these screencasts](#)

	Column	Column 3	Column 4	Column 5
1. Logo der Deutschen Nationalbibliothek				
2. Schriftbanner mit Leipzig, Frankfurt am Main				
3. Katalog der Deutschen Nationalbibliothek				
4. Suchergebnis - Suche nach: "Nationalsozialismus", Treffer 1-500 von insgesamt 13700 - 16.01.19 11:11				
5. Bücher	1			Link zu diesem Datensatz: http://d-nb.info/100900852
6. Titel: Kunst am Pranger : Die Moderne im Nationalsozialismus / Andreas Hüneke				
7. Person(en): Hüneke, Andreas				
8. Ausgabe: 2018. Auflage				
9. Verlag: Paderborn : Fink, Wilhelm				
10. Zeitliche Einordnung: Erscheinungstermin: Dezember 2021				
11. Umfang/Format: 280 Seiten : 40 s/w Abb.				
12. ISBN/Einband/Preis: 978-3-7705-5183-5 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen] Gb. : ca. EUR 34.90 (DE), ca. sfr 49.90 (freier Pr.)				
13. Bestellnummer(n): VLB_ID_old: 3633739				
14. EAN: 9783770551835 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen]				
15. Sprache(n): Deutsch (ger)				edit
16. Beziehungen: Schriften der Berliner Forschungsstelle „Entartete Kunst“				
17. Sachgruppe(n): 700 Künste, Bildende Kunst allgemein				
18. Bücher	2			Link zu diesem Datensatz: http://d-nb.info/117382646



Münzkabinett
 Staatliche Museen zu Berlin

[PATENSCHAFT](#)
[NACHRICHTEN](#)
[KONTAKT](#)
[ÜBER UNS](#)
[DE](#)
[| EN](#)

Interaktiver Katalog des Münzkabinetts

[START](#)
[SUCHE](#)
[KARTE](#)

Erkunden Sie eine der
 größten numismatischen
 Sammlungen der Welt

z. B. Gold Antike Architektur



[ERWEITERTE SUCHE](#)



Schlagwortsuche



Münzstätten und
 Fundorte



Nachrichten

IKMK überschreitet die 33.000



Münzkabinett
Staatliche Museen zu Berlin

PATENSCHAFT NACHRICHTEN KONTAKT ÜBER UNS DE | EN







Interaktiver Katalog des Münzkabinetts

START **SUCHE** KARTE

ZURÜCKSETZEN

1104 TREFFER ANZEIGEN

Antike	Archaik und Klassik (7.-4. Jh. v. Chr.)	Hellenismus (4.-1. Jh. v. Chr.)	Römische Kaiserzeit (1. Jh. v. Chr.-3. Jh.)	Spätantike (3.-5. Jh.)	Mittelalter	Frühmittelalter (5.-9. Jh.)
Hochmittelalter (10.-12. Jh.)	Spätmittelalter (13.-15. Jh.)	Neuzeit	15.-16. Jh.	17.-18. Jh.	19. Jh.	20.-21. Jh.
Berlin und Brandenburg-Preußen	Deutschland	Westeuropa (ohne D)	Osteuropa	Nordeuropa	Italien	Griechenland
Spanien und Portugal	Kleinasien	Vorderer Orient	Asien (ohne Kleinasien und Vord. Orient)	Afrika	Amerika und Australien	Münzstand: Antike Herrscher
Kaiser und Könige	Weltliche Fürsten	Geistliche Fürsten	Städte und Städtebünde	Mythen	Ereignisse	Krieg und Frieden
Porträts	Herrscherrepräsentation	Frauen	Kinder	Götter	Heroen	Heilige
Personifikationen und Allegorien	Berühmte Persönlichkeiten	Christliche Ikonographie	Tiere und Fabelwesen	Pflanzen	Architektur und Stadtansichten	Gegenstände
Heraldik / Wappen	Herstellung und Münztechnik	Gegenstempel, Erasionen u.a.	Fälschungen	Münzschmuck und Schmuckmünzen	Nichtmünzliches und Papiergeld	Medaillen und Modelle

		Aachen: Stadt Jungheit (Aachen) (Deutschland, Rheinland) Groschen, 1374 18206308			Aachen: Stadt Aachen (Deutschland, Rheinland) 1/4 Groschen (Piéfort), 1498 18243445
		Aachen: Stadt Aachen (Deutschland, Rheinland) 1/4 Groschen, 1496 18243447			Alpen: Herrschaft (Deutschland, Rheinland) Pfennig, 1344-1351 18244632
		Alpen: Herrschaft (Deutschland, Rheinland) Heller, 1386-1401 18244633			Alpen: Herrschaft (Deutschland, Rheinland) Heller, 1386-1401 18244634
		Alpen: Herrschaft (Deutschland, Rheinland) Weißpfennig, nach 1436 bzw. 1463 18244635			Alpen: Herrschaft (Deutschland, Rheinland) Schilling, 1422-1465 18244643



Münzkabinett
Staatliche Museen zu Berlin

PATENSCHAFT NACHRICHTEN KONTAKT ÜBER UNS DE | EN

Interaktiver Katalog des Münzkabinetts

START SUCHE KARTE

Tablett



1 / 1104



Aachen: Stadt 1374

Ausstellung im Bode-Museum
Raum 242, BM-054/24 Deutsches Reich (Norden). Silber
13.-15. Jh.

Vorderseite KAROLVS MAG-NVS INPERAT. [Karolus Magnus Imperator]. Brustbild Kaiser Karls des Großen mit Krone, Zepter und Reichsapfel über Adlerschild.

Rückseite XC VINCIT - XC REGNA - AN DNI M- CCCLXXIII / MON-ETA - IVNC-heIT. [Christus Vincit Christus Regnat Anno Domini MCCCCLXXIII Moneta Jungheit]. Langes Kreuz umgeben von doppeltem Schriftkreis.

Dargestellte/r Karl der Große (768-814), König der Franken, seit 800 Kaiser



Münzstand Stadt

Datierung 1374
Spätmittelalter

Nominal Groschen

Material Silber

Herstellung geprägt

Gewicht 2,27 g



Liste

Titel

Pro Seite 100 ▾

Vorderseite

Rückseite

Beide Seiten

Liste

4 5 > »



Aachen: Stadt
Jungheit (Aachen) (Deutschland,
Rheinland)
Groschen, 1374
18206308



Aachen: Stadt
Aachen (Deutschland, Rheinland)
1/4 Groschen (Piéfort), 1498
18243445



Aachen: Stadt
Aachen (Deutschland, Rheinland)
1/4 Groschen, 1496
18243447



Alpen: Herrschaft
(Deutschland, Rheinland)
Pfennig, 1344-1351
18244632



Alpen: Herrschaft
(Deutschland, Rheinland)
Heller, 1386-1401
18244633



Alpen: Herrschaft
(Deutschland, Rheinland)
Heller, 1386-1401
18244634



Alpen: Herrschaft
(Deutschland, Rheinland)
Weißpfennig, nach 1436 bzw. 1463
18244635



Alpen: Herrschaft
(Deutschland, Rheinland)
Schilling, 1422-1465
18244643

OpenRefine_Hands_on_20190116-Hiltmann-Frohmann - Editor

Datei Bearbeiten Format Ansicht ?

Diese Daten wurden dem Interaktiven Katalog des Münzkabinetts der Staatlichen Museen zu Berlin entnommen (<https://ikmk.smb.museum/>) und für die Hands-on Session noch etwas verändert (verunreinigt). Es sind jene Münzen, die mit den Stichworten "Mittelalter", "Deutschland" und "Heraldik / Wappen" versehen sind. (Torsten Hiltmann)

Brandenburg: Johann I. und Otto III. Brandenburg: Johann I. und Otto III.

Brandenburg: Johann I. und Otto III.

(Germany, Brandenburg)

Denar (MA), um 1250

18206726

Waldeck-Pyrmont: GrafschaftWaldeck-Pyrmont: Grafschaft

Waldeck-Pyrmont: Grafschaft

Lügde (Germany, Westfalen)

Denar (MA), 1250-1300

18213795

Brandenburg: MarkgrafschaftBrandenburg: Markgrafschaft

Brandenburg: Markgrafschaft

(Germany, Brandenburg)

Denar (MA), um 1275

18206724

Schlotheim: Herren

Schlotheim: Herren

Schlotheim (Germany, Thüringen)

Brakteat, 1280-1290

18215954

Brandenburg: MarkgrafschaftBrandenburg: Markgrafschaft

Brandenburg: Markgrafschaft

(Germany, Brandenburg)

Denar (MA), um 1290

18206725

Köln: Erzbistum und Waldeck-Pyrmont: GrafschaftKöln: Erzbistum und Waldeck-Pyrmont: Grafschaft

Köln: Erzbistum und Waldeck-Pyrmont: Grafschaft

Lügde (Germany, Westfalen)

Denar (MA), 1297-1304

18213796

Köln: Erzbistum und Waldeck-Pyrmont: GrafschaftKöln: Erzbistum und Waldeck-Pyrmont: Grafschaft

Köln: Erzbistum und Waldeck-Pyrmont: Grafschaft

Lügde (Germany, Westfalen)

Denar (MA), 1297-1304

18213798

Köln: Erzbistum und Waldeck-Pyrmont: GrafschaftKöln: Erzbistum und Waldeck-Pyrmont: Grafschaft

Köln: Erzbistum und Waldeck-Pyrmont: Grafschaft

Lügde (Germany, Westfalen)

Denar (MA), 1297-1304

18213801

Montabaur: StadtMontabaur: Stadt

Daten bearbeiten

127.0.0.1:3333/project?project=1893534342500

Meistbesucht Erste Schritte

OpenRefine Bibl Daten DNB Stichwort Nationalsozialismus 500 Treffer txt [Permalink](#) Open... Export Help

Facet / Filter Undo / Redo 0 / 0

5093 rows Show as: **rows** records Show: 5 10 25 50 rows Extensions: Wikidata

« first » previous 1 - 50 next » last »

▼ All ▼ Column ▼ Column 3 ▼ Column 4 ▼ Column 5

Using facets and filters

Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top of each data column.

Not sure how to get started?
[Watch these screencasts](#)

1.	Logo der Deutschen Nationalbibliothek				
2.	Schriftkatalog mit Leipzig, Frankfurt am Main				
3.	Katalog der Deutschen Nationalbibliothek				
4.	Suchergebnis - Suche nach: "Nationalsozialismus", Treffer 1-500 von insgesamt 13700 - 16.01.19 11:11				
5.	Bücher	1			Link zu diesem Datensatz: http://d-nb.info/100900852 [http://d-nb.info/100900852]
6.	Titel: Kunst am Pranger : Die Moderne im Nationalsozialismus / Andreas Hüneke				
7.	Person(en): Hüneke, Andreas				
8.	Ausgabe: 2018. Auflage				
9.	Verlag: Paderborn : Fink, Wilhelm				
10.	Zeitliche Einordnung: Erscheinungstermin: Dezember 2021				
11.	Umfang/Format: 280 Seiten : 40 s/w Abb.				
12.	ISBN/Einband/Preis: 978-3-7705-5183-5 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen] Gb. : ca. EUR 34.90 (DE), ca. sfr 49.90 (freier Pr.)				
13.	Bestellnummer(n): VLB_ID_old: 3633739				
14.	EAN: 9783770551835 [Titel anhand dieser ISBN in Citavi-Projekt übernehmen]				
15.	Sprache(n): Deutsch (ger)				edit
16.	Beziehungen: Schriften der Berliner Forschungsstelle „Entartete Kunst“				
17.	Sachgruppe(n): 700 Künste, Bildende Kunst allgemein				
18.	Bücher	2			Link zu diesem Datensatz: http://d-nb.info/117382646 [http://d-nb.info/117382646]

Daten bearbeiten

127.0.0.1:3333/project?project=1893534342500

Meistbesucht Erste Schritte

OpenRefine Bibl Daten DNB Stichwort Nationalsozialismus 500 Treffer txt [Permalink](#) Open... Export Help

Facet / Filter Undo / Redo 0 / 0

5093 rows

Show as: rows records Show: 1-50 rows

« first < previous 1 - 50 next > last »

Extensions: Wikidata

Using facets and filters

Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top of each data column.

Not sure how to get started? [Watch these screencasts](#)

Facet

- Facet
- Text filter
- Edit cells
- Edit column
- Transpose
- Sort...
- View
- Reconcile

onalbibliothek
Frankfurt am Main
ationalbibliothek
Transform...
Common transforms
Fill down
Blank down
Split multi-valued cells...
Join multi-valued cells...
Cluster and edit...
Replace

	Column	Column 3	Column 4	Column 5
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				

Link zu diesem Datensatz: <http://d-nb.info/100900852> [http://d-nb.info/100900852]

Link zu diesem Datensatz: <http://d-nb.info/117382646> [http://d-nb.info/117382646]

Daten bearbeiten

OpenRefine OpenRefine_Hands_on_20190116 Hiltmann Frohmann txt [Permalink](#) Open... Export Help

Facet / Filter Undo / Redo 0 / 0

Refresh Reset All Remove All

Column 1 change invert reset

1153 choices total, too many to display
Set choice count limit

Facet by choice counts

2686 records

Show as: rows **records** Show: 5 10 25 50 records

Extensions: Wikidata

« first < previous 1 - 25 next > last »

Column 1

Column 2

1. Facet dem Interaktiven Katalog des MÄnzka...
2. Text filter etwas verändert (verunreinigt). Es sind j...
3. Edit cells und Otto III. Brandenburg. Johann I. und Otto III.
4. Edit column
5. Transpose Transpose cells across columns into rows...
6. Sort... Transpose cells in rows into columns...
7. View Columnize by key/value columns...
8. Reconcile
9. Schlotheim: Herren
10. Brandt, 1280-1290
11. Brandenburg: Markgrafschaft Brandenburg: Markgrafschaft
12. (Germany, Brandenburg)
13. 18206725
14. KÄŕlin: Erzbistum und Waldeck-Pyrmont: Grafschaft
15. Denar (MA), 1297-1304
16. KÄŕlin: Erzbistum und Waldeck-Pyrmont: Grafschaft
17. LÄŕgde (Germany, Westfalen)
18. 18213796
19. KÄŕlin: Erzbistum und Waldeck-Pyrmont: Grafschaft
20. Denar (MA), 1297-1304
21. Montabaur: Stadt Montabaur: Stadt
22. (Germany)
23. 18239334
24. Brandenburg: Markgrafschaft Brandenburg: Markgrafschaft
25. (Germany, Brandenburg)

Daten bearbeiten

127.0.0.1:3333/project?project=2297866406316

Meistbesucht Erste Schritte

OpenRefine OpenRefine_Hands_on_20190116 Hiltmann Frohmann txt Permalink

Open... Export Help

Facet / Filter Undo / Redo 1 / 1

Refresh Reset All Remove All

538 records

Show as: rows records Show: 5 10 25 50 records

« first < previous 1 - 25 next > last »

Column 1 change invert reset

No column named Column 1

Column 1	Column 1 1	Column 1 2	Column 1 3	Column 1 4	Column 1 5	Column 1 6
1. Die Daten wurden dem Interaktiven Katalog des Münzkabinetts der Staatlichen Museen zu Berlin entnommen (https://ikmk.smb.museum/) und für die Hands-on Session noch etwas verändert (verunreinigt). Es sind jene Münzen, die mit den Stichworten "Mittelalter", "Deutschland" und "Heraldik / Wappen" versehen sind. (Torsten Hiltmann)	Brandenburg: Johann I. und Otto III. Brandenburg: Johann I. und Otto III.	(Germany, Brandenburg)	18206726	Waldeck-Pyrmont: Grafschaft		
					Brandenburg: Johann I. und Otto III.	
					Denar (MA), um 1250	
					Waldeck-Pyrmont: GrafschaftWaldeck-Pyrmont: Grafschaft	
					LÄgde (Germany, Westfalen)	
2. Denar (MA), 1250-1300	Brandenburg: MarkgrafschaftBrandenburg: Markgrafschaft	(Germany, Brandenburg)	18206724	Schlotheim: Herren	18213795	
					Brandenburg: Markgrafschaft	
					Denar (MA), um 1275	edit
					Schlotheim: Herren	
					Schlotheim (Germany, Thüringen)	
3. Brakteat, 1280-1290	Brandenburg: MarkgrafschaftBrandenburg: Markgrafschaft	(Germany, Brandenburg)	18206725	KÄlin: Erzbistum und Waldeck-Pyrmont: Grafschaft	18215954	
					Brandenburg: Markgrafschaft	
					Denar (MA), um 1290	
					KÄlin: Erzbistum und Waldeck-Pyrmont: GrafschaftKÄlin: Erzbistum und Waldeck-Pyrmont: Grafschaft	
					LÄgde (Germany, Westfalen)	
4. Denar (MA), 1297-1304	KÄlin: Erzbistum und Waldeck-Pyrmont: GrafschaftKÄlin: Erzbistum und Waldeck-Pyrmont: Grafschaft	LÄgde (Germany, Westfalen)	18213798	KÄlin: Erzbistum und Waldeck-Pyrmont: Grafschaft	18213796	
					KÄlin: Erzbistum und Waldeck-Pyrmont: Grafschaft	
					Denar (MA), 1297-1304	

Daten bearbeiten

OpenRefine OpenRefine_Hands_on_20190116 Hiltmann Frohmann txt [Permalink](#) Open... Export Help

Meistbesucht Erste Schritte

127.0.0.1:3333/project?project=2297866406316

2686 rows

Extensions: Wikidata

Facet / Filter Undo / Redo 1 / 1

Refresh Reset All Remove All

Column 1 change invert reset

No column named Column 1

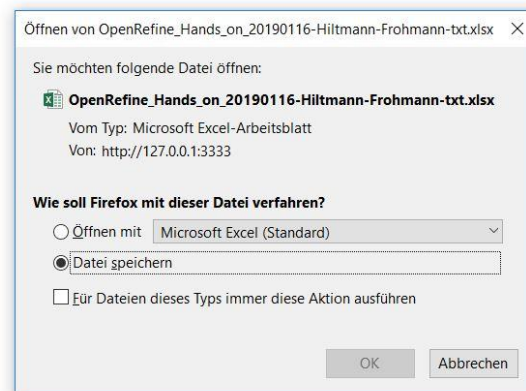
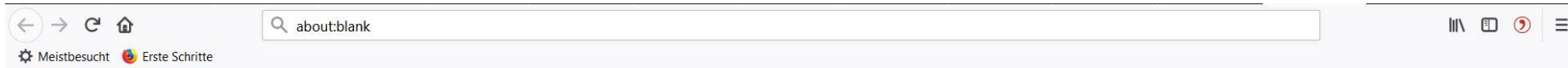
Show as: rows records Show: 3 10 25 50 rows

« first < previous 1 - 25 next > last »

Column 1	Column 1 1	Column 1 2	Column 1 3	Column 1 4	Column 1 5	Column 2
1.	1. Diese Daten wurden dem interaktiven Katalog des MÄnz-kabinetts der Staatlichen Museen zu Berlin entnommen (https://knk.smb.museum/) und fÄr die Hands-on Session noch etwas verÄndert (verunreinigt). Es sind jene MÄnzen, die mit den Stichworten "Mittelalter", "Deutschland" und "Heraldik / Wappen" versehen sind. (Torsten Hiltmann)	Brandenburg: Johann I. und Otto III. Brandenburg: Johann I. und Otto III.	(Germany, Brandenburg)	18206726	Waldeck-Pyrmont: Grafschaft	
2.						Brandenburg: Johann I. und Otto III.
3.						Denar (MA), um 1250
4.						Waldeck-Pyrmont: GrafschaftWaldeck-Pyrmont: Grafschaft
5.						LÄgde (Germany, Westfalen)
6.	Denar (MA), 1250-1300	Brandenburg: MarkgrafschaftBrandenburg: Markgrafschaft	(Germany, Brandenburg)	18206724	Schlotheim: Herren	18213795
7.						Brandenburg: Markgrafschaft
8.						Denar (MA), um 1275 edit
9.						Schlotheim: Herren
10.						Schlotheim (Germany, ThÄringen)
11.	Brakteat, 1280-1290	Brandenburg: MarkgrafschaftBrandenburg: Markgrafschaft	(Germany, Brandenburg)	18206725	KÄn: Erzbistum und Waldeck-Pyrmont: Grafschaft	18215954
12.						Brandenburg: Markgrafschaft
13.						Denar (MA), um 1290
14.						KÄn: Erzbistum und Waldeck-Pyrmont: GrafschaftKÄn: Erzbistum und Waldeck-Pyrmont: Grafschaft
15.						LÄgde (Germany, Westfalen)
16.	Denar (MA), 1297-1304	KÄn: Erzbistum und Waldeck-Pyrmont: GrafschaftKÄn: Erzbistum und Waldeck-Pyrmont: Grafschaft	LÄgde (Germany, Westfalen)	18213798	KÄn: Erzbistum und Waldeck-Pyrmont: Grafschaft	18213796
17.						KÄn: Erzbistum und Waldeck-Pyrmont: Grafschaft

16./23. Januar 2019 | @ubffm | www.ub.uni-frankfurt.de/digitalhumanities | Jakob Frohmann

Daten speichern / exportieren



Nächste Woche gemeinsam...

- Zellen teilen und (wieder) vereinigen mit Hilfe von Separatoren
- Facetieren und Filtern
- Clustern von Daten
- Anreichern von Daten (Beispiele)

Bitte daran denken: Software installieren!

<http://openrefine.org/download.html>

Tipps & Tricks / Links / Literatur

OpenRefine

<http://openrefine.org/>

Ressourcen

<https://www.wikidata.org/> (Daten aller Art, default in OpenRefine)

<https://www.geonames.org/> (Geografika)

<http://swb.bsz-bw.de/DB=2.104/> (Personen, Körperschaften, Konferenzen, Geografika, Sachschlagwörter, Werktitel) [GND - Gemeinsame Normdatei der Deutschen Nationalbibliothek, mehr Infos [hier](#)]

Visualisierung:

<https://geobrowser.de.dariah.eu/>

<http://hdlab.stanford.edu/palladio/>

Tipps & Tricks / Links / Literatur

Blogs

<https://histhub.ch/cat/net/blog/openrefine/> (Blog-Serie zur Arbeit an historischen Daten mit OpenRefine)

<http://blog.lobid.org/2018/08/27/openrefine.html> (einfache Anreicherung von Daten in OpenRefine mit [Personen-] Daten aus der GND via lobid.org)

Literatur

Ruben Verborgh/Max de Wilde, Using OpenRefine. The essential OpenRefine guide that takes you from data analysis and error fixing to linking your dataset to the Web (Community experience distilled), Birmingham, Mumbai 2013. [\[Online-Ressource über UB FFM\]](#)

Danke für Ihre Aufmerksamkeit!

Vielen Dank für die zur Verfügungstellung von Materialien und Daten an Agnes Brauer (Universitätsbibliothek Johann Christian Senckenberg) und Torsten Hiltmann (Zentrum für Digitale Geschichtswissenschaft, Universität Münster).

Workshop konzipiert in Anlehnung an "[Library Carpentry: OpenRefine Lessons for Librarians.](#)" (2016)